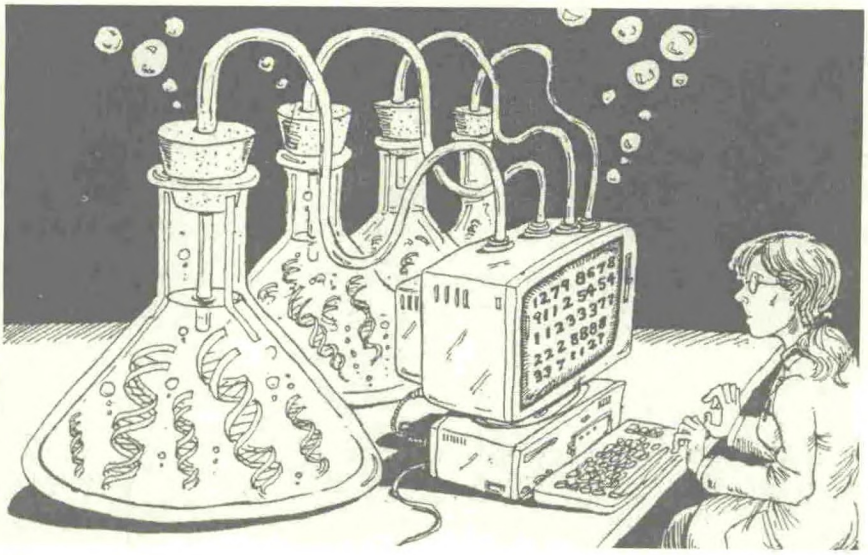
สหัส พรหมสิทธิ์

# DNA คอมพิวเตอร์



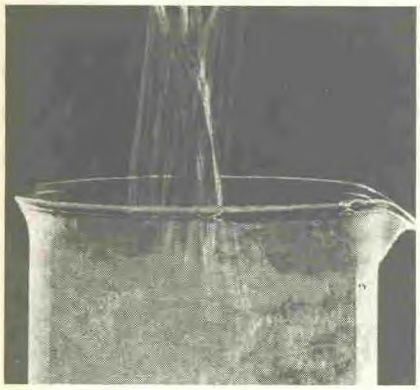
*เทคโนโลยี DNA สามารถที่จะเก็บข้อมูลได้มากกว่าเทคโนโลยีอื่น ๆ ซึ่งแม้แต่สมองของมนุษย์ก็สู้ไม่ได้*

ถ้าจะเปรียบการทำงานของคอมพิวเตอร์ยุคต่อไปที่อาศัย DNA เข้าช่วยในการคำนวณ ก็น่าจะมีบางสิ่งที่เหมือนกับโครงการก่อสร้างขนาดใหญ่มาก ๆ ที่ต้องใช้กำลังคนนับร้อยหรือนับพันเข้าช่วย โดยเฉพาะกับงานก่อสร้างในอดีตทางประวัติศาสตร์ ซึ่งจะมีอภิมหาโครงการให้เห็นอยู่หลายแห่ง นับตั้งแต่การก่อสร้างปิรามิดในประเทศอียิปต์ จนถึงการสร้างกำแพงเมืองจีน ที่ใช้แรงคนเป็นจำนวนนับหมื่น โดยให้แต่ละคนทำหน้าที่เฉพาะอย่าง และเมื่อนำความชำนิชำนาญตลอดจนแรงคนนับเป็นหมื่นมารวมกันเพื่อช่วยกันก่อสร้างหรือแก้ปัญหาขนาดใหญ่ สำหรับโครงการใดโครงการหนึ่ง ก็จะทำให้ปัญหาหรือเรื่องที่แสนจะยากกลับกลายเป็นของง่าย จนสามารถทำได้แล้วเสร็จในเร็ววัน หากได้มีการวางแผนหรือมีวิธีการทำงานที่ดี ดังเช่นการใช้สิ่งมีชีวิตในระดับโมเลกุลมาแก้ปัญหาในทางคำนวณในรูปของ DNA ซึ่งก่อนที่จะพูดถึงการสร้าง**คอมพิวเตอร์ที่ทำงานด้วย DNA** ที่กำลังพัฒนาอยู่ในขณะนี้ ก็จะขอกล่าวถึงความเป็นมาในอดีตเมื่อราว 3 ปีที่แล้ว (1994) ว่ามีใครบ้างที่เป็นผู้ริเริ่มบุกเบิกในด้านนี้

## เทคโนโลยีสารสนเทศใน DNA โมเลกุล

บุคคลที่น่ากล่าวถึงมากที่สุดในฐานะที่พยายามนำสิ่งที่เล็กกว่าทรานซิสเตอร์มาสร้างคอมพิวเตอร์ยุคใหม่ด้วยการอาศัย DNA โมเลกุล ก็คือ **ดร.เลียวนาร์ด เอเดิลแมน** (Dr. Leonard Adleman) ซึ่งเป็นนักวิทยาศาสตร์คอมพิวเตอร์ประจำมหาวิทยาลัยเซาท์เทอร์นแคลิฟอร์เนีย ในสหรัฐฯ ก่อนที่จะพูดถึงรายละเอียดอันเป็นแนวทางในการนำ **DNA โมเลกุล**มาแก้ปัญหาทางคำนวณ ซึ่ง ดร.เอเดิลแมนเป็นผู้ริเริ่ม ก็จะขอทบทวนในสิ่งที่เป็นเสมือนสะพานที่จะเข้ามาเชื่อมโยงโลกของเคมี-ชีววิทยา กับโลกของการคำนวณและการประมวลผลข้อมูล นั่นก็คือความต้องการขั้นพื้นฐานสำหรับเทคโนโลยีสารสนเทศขั้นสูง (Advanced IT) ก็คือ การมีสื่อในการเก็บข้อมูลที่มี **Memory-location** เป็นจำนวนมาก โดยแต่ละ Location จะต้องสามารถที่จะอยู่ในสถานะที่เรียกว่า a discrete number of states เช่นในรูปของดิจิตอล

ตัวอย่างของเรื่องนี้สำหรับในโลกของสิ่งที่มีชีวิตก็จะเป็น**เทคโนโลยีสารสนเทศ** (IT) ของ genes ที่จะอยู่ในรูปของดิจิตอลเช่นกัน ความจริงอันนี้ได้มีการค้นพบมานานแล้วด้วยการค้นคว้าของ **เมนเดล** (Gregor Mendel) ที่ว่า ในการสืบพันธุ์ทางชีววิทยานั้นจะไม่มีการสืบทอดพันธุกรรมแบบผสมผสานโดยเอาคุณสมบัติต่างมาจากของทั้งพ่อและแม่แบบรวมกันแล้วหารด้วยสอง แต่จะเป็นไปในรูปของสืบทอดแบบ **discrete particles** ชนิดจะต้องเลือกเอาอย่างใดอย่างหนึ่งแบบ YES หรือ NO เท่านั้น



จากนี้ก็จะมาถึงจุดที่ว่าเนื้อหาอันเป็นสื่อที่อยู่ในเมนสตอเรจของเซลล์ของสิ่งมีชีวิตจะอยู่ในรูปที่เป็นวัตถุธาตุทางเคมี แทนที่จะอยู่ในสภาพที่เป็น**อิเล็กทรอนิกส์** นอกจากนี้ก็ยังมีความจริงอยู่ข้อหนึ่งว่า โมเลกุลบางชนิดมีคุณสมบัติในแง่ของ polymerizing ในแง่ของการสามารถเรียงตัวกันเป็นลูกโซ่ขนาดยาวที่มีความยาวอันไม่จำกัด ในโลกของเรานี้ได้มี**โพลีเมอร์** (polymer) อยู่หลายชนิด เช่น polymerized ethylene ที่เกิดจากการเรียงตัวของโมเลกุลเล็ก ๆ ที่เรียกว่า ethylene หรือโพลีเมอร์ของน้ำตาลที่เกิดจากการรวมตัวของแป้งกับเซลลูโลส เมื่อก้าวไปอีกขั้นก็จะเห็นว่าโพลีเมอร์บางชนิดนั้นแทนที่จะมีสภาพเป็นลูกโซ่ที่ประกอบด้วยโมเลกุลเล็กเช่น ethylene แต่เพียงอย่างเดียว กลับปรากฏว่าเป็นลูกโซ่หรือเชน (chains) ที่ประกอบด้วยโมเลกุลตั้งแต่สองชนิดขึ้นไปที่แตกต่างกัน และทันทีที่ความหลากหลาย (heterogeneity) ได้ก้าวเข้าสู่โลก**ของโพลีเมอร์เชน** ก็จะทำให้การเข้ามาเกี่ยวข้องของ IT เกิดความเป็นไปได้ในทางทฤษฎี เช่น ถ้ามีโมเลกุลอยู่**สอง**ชนิดในเชน (chain) ก็สามารถนำมาแปลเป็นความหมายในแง่ของ digital information ได้ว่า โมเลกุลแต่ละชนิดจะแทนเลข 1 และ 0 ตามลำดับ โดยจะสามารถเก็บข้อมูลที่เป็นอะไรก็ได้เอาไว้ในเชนดังกล่าว ถ้าหากว่ามีความยาวอย่างเพียงพอ

โพลีเมอร์ที่สำคัญอีกชนิดที่อยู่ในเซลล์ของสิ่งมีชีวิตก็คือพวก polynucleotides โดยแบ่งได้ออกเป็นสองพวกใหญ่ ๆ นั่นก็คือ DNA กับ RNA ทั้งสองจะเป็นเชนของโมเลกุลขนาดเล็กที่เรียกว่า nucleotides ภายในเชนที่ประกอบไปด้วยความหลากหลายเช่น DNA กับ RNA นี้จะประกอบด้วยเชนต่าง ๆ ที่มี nucleotide อยู่ 4 ชนิด อันเป็นโอกาสที่เทคโนโลยีของการเก็บข้อมูลจะก้าวเข้ามามีบทบาท คือแทนที่จะมีอยู่สองสถานะคือ 1 กับ 0 IT ของเซลล์ของสิ่งมีชีวิตจะมีอยู่ 4 สถานะด้วยกัน โดยมีชื่อว่า A, T, C และ G ซึ่งในบางหลักการนั้นถือว่ามีข้อแตกต่างกันน้อยมาก ระหว่าง IT ที่มีสองสถานะ (two-state binary) ที่มนุษย์นำมาใช้ในคอมพิวเตอร์ที่ทำงานด้วยทรานซิสเตอร์ กับ IT ที่มีสถานะของเซลล์มีชีวิต ซึ่งใน**แต่ละเซลล์**ของสิ่งมีชีวิตเช่นของมนุษย์ก็จะสามารถเก็บข้อมูลที่เป็น Encyclopaedia Britannica ได้ถึง 30 เล่ม

จากนี้ก็จะวกกลับมาพูดถึงการค้นคว้าของ ดร.เอเดิลแมนต่อไป โดยจะเริ่มที่การใช้ DNA โมเลกุลในการแก้ปัญหาง่าย ๆ เป็นลำดับแรก

## การค้นคว้าของ ดร.เอเดิลแมน

ในตอนแรกเริ่มของการค้นคว้าของ ดร.เอเดิลแมนได้มีการใช้ DNA โมเลกุลจำนวนหนึ่งเพื่อที่จะแก้ปัญหาทางคำนวณซึ่งเป็นที่รู้จักกันดี นั่นก็คือเรื่องที่เกี่ยวกับ travelling salesman ที่ปรากฏอยู่ในวิชาวิจัยเชิงปฏิบัติการ (operations research) ขนาดของปัญหาที่ ดร.เอเดิลแมนใช้สำหรับในขั้นต้นนี้จะมีเมืองต่าง ๆ อยู่ 7 เมืองที่เชื่อมโยงด้วยถนนแบบวันเวย์ จำนวน 14 สาย ปัญหาก็คือว่าให้หาเส้นทางเพียงเส้นเดียวที่จะทำให้พนักงานขาย (salesman) เดินทางจากเมืองหนึ่งอันเป็นจุดเริ่มต้นและไปถึงจุดหมายปลายทางโดยผ่านเมืองทุกเมืองเพียง**ครั้งเดียว**เท่านั้น วิธีแก้ปัญหาของ ดร.เอเดิลแมนสำหรับในเรื่องนี้ก็คือการสร้างแถบหรือสแตรนด์ (strand) ของ DNA ที่มีขนาดสั้น ๆ ขึ้นมาเป็นจำนวน 7 แถบด้วยกันเพื่อแทนเมืองต่าง ๆ โดยในแต่ละแถบของ DNA จะมีเบส (base) ต่าง ๆ อยู่ 4 ชนิดคือ A, C, T และ G โดยที่ชื่อต่าง ๆ ของเมืองทั้งหมดจะถูกเรียงกันเป็นแถวที่มีความยาวเท่ากับ 20 เบส จากนี้ ดร.เอเดิลแมนก็จะสร้าง sequence ต่าง ๆ ที่สัมพันธ์กับถนน 14 สายที่เชื่อมโยงเมืองต่าง ๆ การออกแบบ sequence ต่าง ๆ ที่ทำหน้าที่แทนเส้นทาง (routes) นั้นจะเป็นไปในทำนองที่พ่วงโยงทีละสองเมืองซึ่งถูกต่อเชื่อมด้วยเส้นทางที่กำหนดไว้ ฯลฯ

ขั้นต่อไปก็ได้มีการนำ sequence ต่าง ๆ เหล่านี้มาคลุกเคล้ากันในหลอดทดลอง (test tube) ซึ่งมันก็จะจับติดกันด้วยรูปแบบที่นับไม่ถ้วน อันเนื่องมาจากการถ่ายทับกันระหว่าง sequence ต่าง ๆ ที่แทนถนนและ sequence ต่าง ๆ ที่ได้รับการสมมติให้แทนเมืองทั้ง 7 ขั้นต่อไปก็เป็นการอาศัยเทคนิคจากวิชาชีววิทยาของโมเลกุล (molecular biology) มาแยก DNA โมเลกุลที่มี sequence ของเมืองต่าง ๆ ที่มีความถูกต้องเป็นไปได้ออกมาจาก sequence ที่เกาะกันเป็นกลุ่ม sequence ที่มีการเรียงลำดับของเมืองอย่างถูกต้องนี้จะต้องมีความยาวตั้งแต่ต้นจนจบที่มีเมืองอยู่ 7 เมือง ซึ่งเมื่อนำไปแปลความหมายด้วยกรรมวิธีทางชีววิทยาของโมเลกุลก็จะให้คำตอบในการแก้ปัญหาเกี่ยวกับ travelling salesman ส่วนโมเลกุลอื่น ๆ อันนับเป็นล้าน ๆ ตัวที่ไม่ต้องการก็จะต้องทิ้งไป

การแก้ปัญหาที่กล่าวในข้างต้นนี้สามารถทำได้โดยอาศัยเพียงแค่กระดาษกับดินสอ แต่เหตุที่นักวิทยาศาสตร์ให้ความสนใจในอันที่จะนำ DNA computing มาแก้ปัญหาชนิดนี้ก็เพราะได้เล็งเห็นว่า การคำนวณที่อาศัย DNA นี้จะมีอานุภาพมากกว่าวิธีอื่นเมื่อต้องการแก้ปัญหาขนาดใหญ่มาก ๆ อันเนื่องจากว่า DNA โมเลกุลนับล้าน ๆ ตัวที่อยู่ในหลอดทดลอง จะทำการคำนวณอย่างพร้อมกันหมดในขณะที่ sequence ต่าง ๆ ได้จับตัวกันเป็นปฏิกิริยาทางชีววิทยาทางด้านโมเลกุล งานค้นคว้าทดลองของ ดร.เอเดิลแมนเท่าที่พูดมานี้ได้เป็นที่สนใจของนักวิทยาศาสตร์หลายท่าน และหนึ่งในจำนวนนี้ก็คือ **ดร.ริชชาร์ด ลิปตัน** (Richard Lipton) ซึ่งเป็นอาจารย์มหาวิทยาลัยปรินซ์ตัน ดร.ลิปตันได้แสดงให้โลกรู้ว่าวิธีการ

ของ ดร.เอเดิลแมนนั้นสามารถที่จะนำมาใช้แก้ปัญหาทางตรรกศาสตร์ที่เรียกว่า Boolean satisfactions problems ได้ ลักษณะของปัญหานี้จะเป็นการพิสูจน์ว่าจะต้องมีเงื่อนไขอะไรบ้างที่จะทำให้กระบวนการทางตรรกวิทยาชุดใดชุดหนึ่งเกิดความเป็นจริง

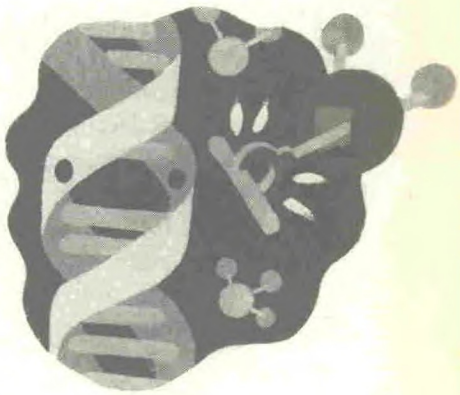
ข้อเสนอในการแก้ปัญหาด้านตรรกศาสตร์ของ ดร.ลิปตันได้เป็นข้อเสนอในทางทฤษฎีว่า statement ต่าง ๆ ที่อยู่ใน propositions นั้นจะเป็นเสมือนเมืองต่าง ๆ ที่ ดร.เอเดิลแมนใช้แก้ปัญหา และความสัมพันธ์ระหว่าง statement ต่าง ๆ ก็จะเปรียบเหมือนถนนหรือเส้นทางที่ใช้เชื่อมโยงเมืองต่าง ๆ ในเรื่องของ travelling salesman ปัญหาทางตรรกศาสตร์ที่ ดร.ลิปตันเป็นผู้เสนอขึ้นมานี้ได้เป็นที่สนใจของ ดร.เอเดิลแมน ซึ่งได้เป็นผู้ให้คำแนะนำว่าถ้าจะแก้ปัญหานี้ด้วย DNA คอมพิวเตอร์ก็จะต้องอาศัยตัวแปร **(variables)** เป็นจำนวน 70 ตัว และสิ่งที่จะมาแทนตัวแปรดังกล่าวก็คือ sequence ต่าง ๆ ของ DNA โดยแต่ละ sequence จะมีความยาวเท่ากับ 720 ฐาน (720 bases long) ความซับซ้อนของ DNA sequence ในกรณีนี้จะเกิดจากจำนวนตัวแปรค่าทั้ง 70 ตัวที่แต่ละตัวจะทำหน้าที่เป็น sta- tement อันหนึ่งที่ประกอบด้วยสองสถานะคือถูก (true) กับผิด (false) ซึ่งเมื่อนำเลขสองคูณตัวของมันเอง 70 ครั้ง โอกาสของความเป็นไปได้ต่าง ๆ จึงเท่ากับ $2^{70}$ การแยก DNA โมเลกุลชุดใดชุดหนึ่งออกมาจากสมาชิกอื่น ๆ นับล้าน ๆ จึงต้องใช้เวลาพอสมควร ซึ่งอาจต้องใช้เวลามากถึง 1,000 วินาที เมื่อเทียบกับการใช้เวลาแค่หนึ่งส่วนพันล้านของวินาทีเมื่อทำการคำนวณด้วยอิเล็กทรอนิกส์ซูเปอร์คอมพิวเตอร์ที่เร็วมาก ๆ

## จากอดีตสู่ปัจจุบัน

การแก้ปัญหาด้วย DNA คอมพิวเตอร์สำหรับในขณะนี้ยังถือว่าต้องใช้เวลาค่อนข้างมาก เพราะมีจำนวนโมเลกุลให้ต้องแยกแยะเป็นจำนวนมากในระดับล้าน ๆ (trillious of molecules) อันทำให้ต้องค้นหาวิธีการในการคำนวณที่ดีกว่าของ ดร.เอเดิลแมนว่าทำอย่างไรจึง**ไม่**ต้องใช้ DNA โมเลกุลเป็นจำนวนมาก ๆ เพื่อให้การแก้ปัญหาเช่นนี้เหมาะสำหรับทางปฏิบัติในห้องทดลอง

การระดมความคิดในเรื่องนี้ได้เป็นที่มาของการประชุมทางวิชาการเป็นครั้งแรกเมื่อเดือนเมษายน 1995 โดยมี ดร.ลิปตัน และ ดร.อีริก บอม (Eric Baum) เป็นผู้จัด ซึ่งทั้งสองท่านนี้ก็เป็นนักวิจัยอยู่ที่ **NEC Research Centre** ในปรินซ์ตัน การประชุมครั้งนั้นได้มีนักวิจัยจากที่ต่าง ๆ เข้ามาร่วมประชุมถึง 200 คน ซึ่งทุกคนก็ได้วาดฝันไว้ว่าการสามารถสร้างคอมพิวเตอร์ที่มีพลังในการคำนวณอันมหาศาลนั้น จะทำให้สามารถนำคอมพิวเตอร์เช่นนี้ไปแก้ปัญหาขนาดใหญ่แค่ไหนก็ได้ จนเป็นผลให้มีการพัฒนาแนวคิดต่าง ๆ ที่ใช้ในด้าน**การคำนวณด้วย DNA** ไปไกลพอสมควรนับจนถึงขณะนี้ ข้อคิดอันหนึ่งที่เห็นว่าน่าสนใจก็คือจินตนาการของ ดร.บอมที่เกี่ยวกับการนำ**DNA เทคโนโลยี**มาสร้างฐานข้อมูล คือภายในฐานข้อมูล DNA นั้นจะสามารถบรรจุคำต่าง ๆ ที่มากจนเหลือคณานับ (a hundred thousand trillion words) ซึ่งแต่ละคำก็จะประกอบด้วยจำนวนบิต (bits) นับเป็นพัน เทคโนโลยีเช่นนี้จึงสามารถที่จะเก็บข้อมูลได้มากกว่าเทคโนโลยีอื่น ๆ ซึ่งแม้แต่สมองของมนุษย์ก็สู้ไม่ได้ อย่างไรก็ตาม การแก้ปัญหาเช่นนี้ด้วยกรรมวิธีของ ดร.เอเดิลแมนก็จะต้องใช้ DNA นับเป็นตัน อันทำให้ต้องหาวิธีคำนวณที่รัดกุมด้วยการทำวิจัยกันอย่างขะมักเขม้นตั้งแต่อดีตเมื่อราวสองปีมาแล้วจนกระทั่งปัจจุบัน

งานค้นคว้าในด้านการคำนวณด้วย DNA ที่เป็นข่าวเมื่อเร็ว ๆ นี้ (สิงหาคม 1996) ก็มี อาทิเช่น กรณีของนาย**มิตซูนอริ โอกิฮาร่า** (Mitsunori Ogihara) แห่งมหาวิทยาลัยโรเชสเตอร์ ในสหรัฐฯ ซึ่งเชื่อว่าเขาได้พบอัลกอริธึม (algorithm) ในทางคำนวณที่สามารถ**ลดขั้นตอน**ต่าง ๆ ในการจัดการกับ DNA โมเลกุลให้ลงได้มาก โดยอาศัยกรรมวิธีที่เรียกว่า**แพตเทอร์นแมตชิง** (pattern matching) จากวิชา AI



การทำงานของอัลกอริธึมแบบใหม่ของโอกิฮาร่าจะเป็นไปในรูปที่ว่า มันจะเข้าไปชะงักการเติบโตของ families ของ strand ต่าง ๆ ที่เห็นว่าไม่มีประโยชน์สำหรับการแก้ปัญหาอย่างใดอย่างหนึ่ง และถ้าทำได้สำเร็จตามนี้ก็จะได้ผลของการค้นคว้าในด้าน DNA คอมพิวเตอร์ที่รุดหน้าไปอย่างมากว่า DNA เพียงหนึ่งปอนด์ จะสามารถแก้ปัญหาต่าง ๆ ได้มากกว่าการนำซูเปอร์คอมพิวเตอร์ทั่วทั้งโลกมารวมกัน ซึ่งโอกิฮาร่าก็จะพิสูจน์แนวความคิดของเขาในก่อนสิ้นปีนี้ (1996) โดยอาศัย**DNA คอมพิวเตอร์ที่เป็นแบบจำลอง** (simulated DNA computer)

## สรุป

ทัศนคติของนักชีววิทยากลุ่มหนึ่งในปัจจุบันได้เป็นไปในทำนองที่ว่าชีววิทยานั้นน่าจะจัดเป็น information science อย่างหนึ่ง ความเห็นเช่นสามารถพิสูจน์ได้ด้วยการหาวิธีการใหม่ ๆ ในการประมวลผลข้อมูลโดยอาศัยโมเลกุลของสิ่งมีชีวิต ซึ่งถ้าใครทำได้สำเร็จก็จะสามารถเก็บเกี่ยวผลประโยชน์ในโลกของการคำนวณได้อย่างมากมาย ความเป็นไปได้ในเรื่องนี้มีสูงมาก เพราะธรรมชาติได้ทำการประมวลผลข้อมูลด้วย DNA มานานนับล้านปี และก็เป็นที่แน่นอนว่า ธรรมชาติอาจไม่รู้วิธีที่ดีที่สุดก็เป็นได้